

文章编号:0253 - 4339(2018) 03 - 0001 - 06
doi:10. 3969/j. issn. 0253 - 4339. 2018. 03. 001

基于数据挖掘技术的地铁站环控系统用能诊断

刘佳慧¹ 龙 静² 潘志刚² 陈焕新¹ 刘江岩¹ 黄荣庚¹ 李正飞¹

(1 华中科技大学制冷及低温工程系 武汉 430074; 2 广州市地下铁道总公司 广州 510030)

摘 要 本文提出了一种地铁站环控系统用能诊断方法,通过数据挖掘技术建立评价模型来评价地铁环控系统的用能特性。首先,通过相关性分析,确定影响地铁环控系统能耗的关键变量:室外温度、客流量;其次,根据所选取的关键变量,采用决策树划分不同的用能模式,进而根据各个模式建立相应的的用能基准;最后,根据不同模式的用能基准对地铁环控系统实际运行数据进行用能诊断。诊断结果表明:环境和客流的变化会引起用能水平的波动,但仍然贴近用能基准。该用能诊断方法能够诊断地铁站用能、识别异常用能模式和识别低能耗的用能模式,有助于地铁站运营漏洞、故障排查和环控模式优化,为单个地铁站的节能工作提供理论依据和实际参考。

关键词 数据挖掘;地铁环控系统;用能诊断;能耗等级;节能优化

中图分类号:TU831.3; TP306.3; U231+.92

文献标识码:A

Energy Consumption Diagnosis Method based on Data Mining Technology in HVAC System in Subway Stations

Liu Jiahui¹ Long Jing² Pan Zhigang² Chen Huanxin¹ Liu Jiangyan¹
Huang Ronggeng¹ Li Zhengfei¹

(1. Department of Refrigeration & Cryogenics, Huazhong University of Science and Technology, Wuhan, 430074, China; 2. Guangzhou Metro Corporation, Guangzhou, 510030, China)

Abstract This paper proposes an energy consumption diagnosis method and establishes an evaluation model based on data mining technology to evaluate the condition of energy utilization of HVAC system in subway stations. Firstly, through the correlation analysis, the key variables that influence the energy consumption of the subway environmental control system are determined: outdoor temperature and passenger flows. Then, according to the selected key variables, different energy utilization modes are generated by using the decision tree and after that the energy benchmark is established on the basis of the energy consumption models. Finally, an energy consumption diagnosis method based on the energy benchmark of different energy utilization models is used in actual data of HVAC system. The results of energy consumption diagnosis show that energy levels can be influenced by the change of environment and passenger flow but still conform to energy benchmark. This energy consumption diagnosis method can diagnose the condition of energy utilization and recognize models of abnormal or low energy consumption, which can optimize energy utilization and provide a theoretical basis and practical reference value for the energy saving a of single subway station.

Keywords data mining; HVAC system in subway stations; energy consumption diagnosis; energy rating; optimization and energy saving

近年来我国大力发展城市轨道交通,截至 2016 年,国内已有 29 座城市开通轨道交通运营线路,运营线路达 130 条,总里程达 3 849 公里。地铁作为城市轨道交通之一,能耗巨大,是城市能源消耗的重要组成部分,地铁系统中供热通风和空调系统(环控系统)的能耗占比较大^[1]。Wang Yongcai 等^[2-3]分析北京地铁站环控系统能耗,指出北京地铁仅 2012 年

环控系统在夏季每月耗能可达 $1.819 \times 10^7 \text{ kW} \cdot \text{h}$ 。Zhu Yingxin 等^[4-5]指出在地铁运行期间,环控系统能耗与列车的牵引耗能相当。地铁环控系统能耗巨大,需尽快开展地铁站节能工作。为了高效、有针对性地开展节能工作,首先需要理清什么是节能。更进一步,需思考该以怎样的基准值评价地铁站是否节能^[6]。建立有效的地铁耗能指标评价体系,能够科

基金项目:国家自然科学基金(51576074,51328602)资助项目。(The project was supported by the National Natural Science Foundation of China (No. 51576074 & No. 51328602).)

收稿日期:2017 年 6 月 9 日

学、公正地评价地铁系统的用能水平,识别用能潜力,这为地铁站的设计、规划、运行以及节能工作的深入开展提供理论依据及实际指导。

专家学者们在轨道交通建设的能耗指标体系有较为广泛的研究。石静雅等^[7]研究了地铁线路类型、子系统类型、专业类型、客流状价等层次以树状方式建立能耗指标体系。孙宁等^[8]根据设备监理实践,完善了城市轨道交通能耗指标体系,归纳了主要设备系统节能措施。郑玲芳^[9]分析各线路牵引能耗和动力能耗,综合考虑影响能耗的因素,构建了适合于轨道交通网络的能耗指标体系,综合地铁多个子系统和影响能耗因素,根据能耗水平和权重等建立了地铁能耗评价体系。已有文献主要针对不同地铁站的用能评价,但对于单个地铁站,其能耗受客流、列车、室外气候等多个因素影响,用能模式多样,已有文献中缺乏对此方面的研究。

本文提出了一种基于数据挖掘技术^[10]的地铁环控系统能耗诊断体系,根据不同用能模式建立能耗基准来评价地铁环控系统的用能特性,为单个地铁站的节能运行、节能改造工作提供理论依据和实际参考。

1 决策树原理

本文采用数据挖掘中的决策树来建模划分地铁站用能模式,进行地铁站用能诊断。决策树是数据挖掘技术中被广泛应用的分类算法,主要通过一系列规则对数据进行分类,提供一种在什么条件下会得到什么值的类似规则的方法,算法主要包括 ID3^[11-12]、C4.5^[13]、CART^[14]、SLIQ^[15]等。决策树主要分为分类树和回归树两种,分类树对离散变量做决策树,回归树对连续变量做决策树。本文采用条件推理树 ctree(conditional inference tree)作为决策树模型建立的算法。ctree 算法主要针对连续变量,其决策树模型构建步骤如下:

1) 选择一个属性放置在根节点,为每个可能的属性值产生一个分支;2) 将样本划分成多个子集,一个子集对应于一个分支;3) 在每个分支上递归地重复此过程,仅使用真正到达该分支的样本;4) 如果在一个节点上的所有样本拥有相同的类别,即停止该部分树的扩展构造决策树(集合划分)时选择属性。

2 数据采集与关键变量选择

本文采集北方某城市地铁站环控系统的实际运行数据进行地铁站用能诊断。实地采集北方该城市某一地铁站在 2013 年夏季 8 月下旬 4 d 的实际运行数据,通过布置于地铁站内外的温度、湿度、CO₂ 传感

器等按分钟采集地铁站环境变量(地铁站内外环境温度、相对湿度及 CO₂ 浓度等),采用电表、传感器等监测得到地铁环控系统中各设备(2 个送风机、2 个排风机、2 个回风机、4 台冷却泵、4 台冷冻泵、3 台冷机及一系列阀等)的运行参数和能耗,通过地铁运营公司采集得到该地铁站每小时的发车对数、客流量(每小时进出站人数之和)。对每小时的发车对数、客流量线性插值得 5 min 内的发车对数和客流量,同时计算得 5 min 内室内外环境的平均温度、相对湿度及 CO₂ 浓度等,最后筛选出每列变量(5 min 内的室内外环境的温度、相对湿度、CO₂ 浓度、发车对数、客流量、环控系统总能耗等)937 个数据样本。

建立决策树模型之前,需要对变量进行相关性分析,去除冗余变量,选出影响地铁能耗的关键变量。根据建立的数据集,选择多个变量:室外环境温度、室外环境相对湿度、发车对数、客流量等。

本文采用圆型矩阵图,可视化展示变量间相关系数,如图 1 所示。为了便于显示图像信息,截取变量关系矩阵图的上半区、下半区为变量间的相关系数。对角线上的圆半径最大,颜色最深,表示变量与自身的相关系数为 1。圆的半径越大、颜色越深,表明圆形所对应两个变量的线性相关系数越大。以图 1 系统能耗行中第 2 个圆为例,表明系统能耗与客流量变量之间的相关性,相关系数为 0.55。

由图 1 可知,客流量、发车对数、室外环境温度等变量与能耗之间有较强的正相关性,表明客流量越大,发车对数越多,室外环境温度越高,地铁环控系统能耗越高。室外环境相对湿度与地铁站能耗存在负

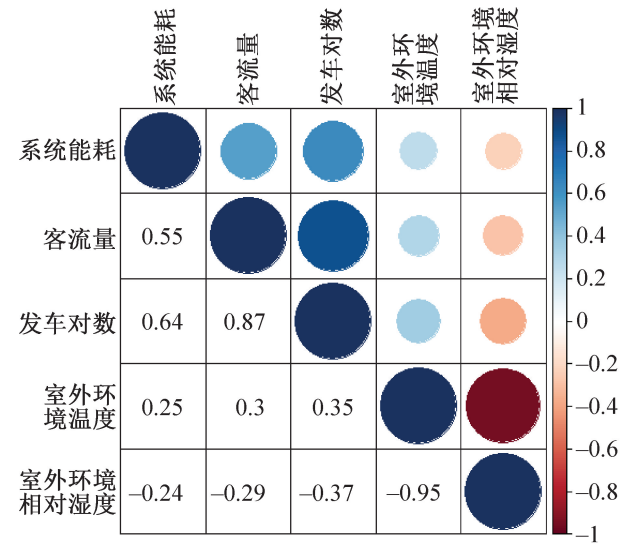


图 1 自变量与因变量之间的相关性分析结果
Fig. 1 Correlation analysis results between the dependent and independent variables

相关性,室外环境相对湿度越低,地铁环控系统用能越高。同时,发车对数与客流量有 0.87 的正相关性,客流量越大,发车对数越多,而客流量与环控系统能耗成正比,因此发车对数与系统能耗成正相关。客流量与系统能耗有因果关系,而发车对数的多少与环控系统能耗并无直接关系,发车对数主要直接影响地铁牵引能耗。最后,选择自变量室外环境温度、客流量作为关键变量,因变量为系统能耗。

3 结果与讨论

3.1 划分用能模式

通过构建条件推理树来分析所选变量与地铁环控系统能耗之间的潜在关系。图 2 为根据所选变量构建的 ctree 决策树模型。系统能耗作为决策树根节点的输出变量,而所选的室外温度和客流量用于构建决策树的树结构。节点 1 选用客流量作为分类属性,将数据分为高客流量和低客流量两类,表明该节点处客流量对能耗数据影响更大。节点 2 和节点 9 选用室外温度作为分类属性,将数据分为高温和低温两类,表明这两节点处室外温度对能耗数据影响更大。节点 3 选用客流量作为分类属性,将低客流量、低室外温度的数据继续分为客流量不同的两类数据,该节点处客流量对能耗数据

影响更大。节点 4 选用室外温度作为分类属性,将较低客流量、低室外温度的数据分成室外温度不同的两类数据,该节点处室外温度对能耗数据影响更大。节点 11 选用客流量作为分类属性,将高客流量、高室外温度的数据分为客流量不同的两类数据,说明该节点处客流量对能耗数据影响更大。图 2 中环控系统能耗被划分成 7 个叶子节点分区,根节点处的箱型图表明各叶子节点处环控系统能耗的分布。将能耗分布相近的区间划分为同一用能模式,可得到 4 个不同的用能模式。

根据决策树模型划分不同的用能模式,根据各个模式中能耗中位值作为用能基准。表 1 为该地铁环控系统的不同用能模式的用能基准。地铁站不同用能模式的特点:模式 1,客流量小,室外环境温度低,环控系统能耗极低,用能模式主要发生在早晨与深夜;模式 2,室外温度偏低,位于某范围内,客流量较小,系统能耗偏低,主要发生在上午客流高峰期前和下午客流量高峰期后;模式 3,客流量较小,环境温度升高,系统能耗增加,用能模式发生于上下午客流高峰期之间、室外环境温度较高的时期;模式 4,客流量增加,而室外环境温度无明显影响,地铁环控系统能耗较高,用能模式发生于上午和下午客流高峰时期。

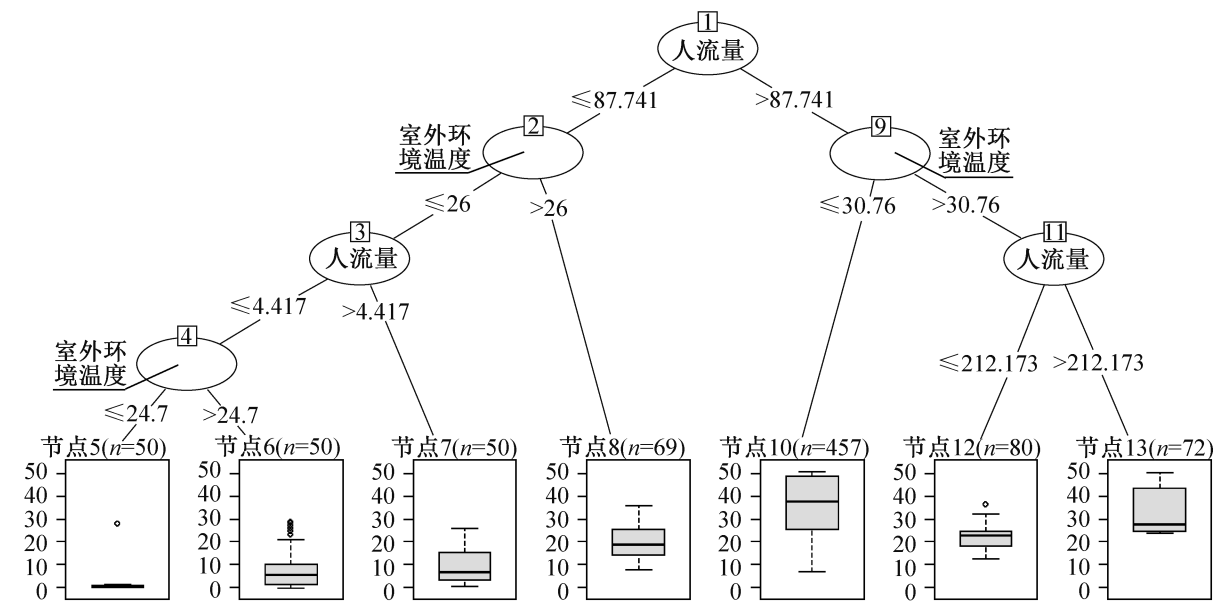


图 2 决策树模型

Fig. 2 Decision tree model

根据不同用能模式特点可知,客流量是影响地铁环控系统能耗的关键因素。系统能耗随客流量增加而提高。当乘客从站外进入地铁,会带入高于站内温度的热空气;当乘客离开地铁时,站内冷空气会随人

员流动带出地铁站,大量的冷空气流向站外,导致站内冷量不足。夏季乘客自身温度较高,客流量增加后,滞留在站台的乘客发热量增大,地铁站内温度提高。综上所述,客流量增加使站内冷量不足而温度升

高,环控系统需提高运行功率来冷却站内空气,降温调整舒适度,因此环控系统能耗增加。

分析室外环境温度与地铁站用能的相关性,室外环境温度为影响地铁系统能耗的关键因素。地铁站进出口多,人员流动大,站外大量新风进入地铁站,室外环境温度越高,进入站内的新风温度越高,站内温度升高。此时环控系统需要消耗更多的电能来维持站内温度和舒适度,能耗升高。

表 1 该地铁环控系统的不同用能模式的用能基准

Tab.1 Energy reference of different models in HVAC in the subway station

模式	节点	客流量 p	室外温 度/℃	用能基准/ (kW·h)
1	5	$p \leq 4.42$	$t \leq 24.7$	0.06
2	6、7	$p \leq 87.74$	$24.7 < t \leq 26$	5.66
3	8、12	$p \leq 212.173$	$t > 26$	21.89
4	10、13	$p > 212.173$	—	35.63

3.2 建立用能基准

根据不同模式的用能基准,按照式(1)计算地铁站环控系统的能耗等级,来评价地铁环控系统的用能特性。

$$R = \frac{E_r}{E_m} \times 100$$

(1)

式中:R 为计算所得能耗等级; E_r 为该地铁站环控系统的实际能源消耗值; E_m 为地铁环控系统不同模式的用能基准。若地铁环控系统实际能耗与同模式下的用能基准相同,则能耗等级为 100;若实际能耗高于用能基准,则能耗等级大于 100,表明地铁站用能高,有潜在的节能空间;若实际地铁站能耗低于用能基准,则能耗等级低于 100,表明地铁站用能低,节能运行。

如图 3 所示,将 H. S. Park 等^[16]划分写字楼的能耗等级的方法应用于地铁环控系统用能诊断研究中,将地铁站环控系统能耗等级分为 A(0 ~ 55)、B(56 ~ 85)、C(86 ~ 115)、D(116 ~ 145)、E(> 145)。由于等级 100 在 C 内,为评价 C 能耗等级内的能耗高低,将 C 分为低于 100 的能耗等级 C1(86 ~ 100)和高于 100 的能耗等级 C2(101 ~ 115)。A、B、C1 表示低于 100 的能耗等级,地铁站环控系统用能低。C2、D、E 表示高于 100 的能耗等级,地铁站环控系统用能高,表明系统有巨大的节能空间。

3.3 模型验证

根据能耗等级系统对地铁站实际数据进行用能

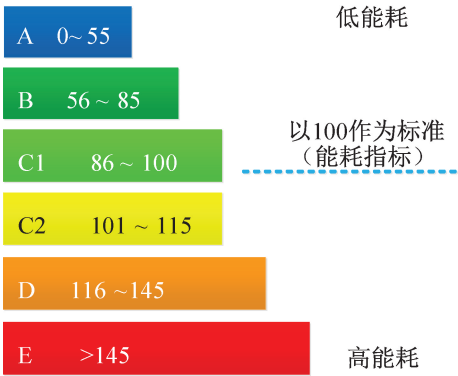


图 3 地铁环控系统能耗等级系统

Fig.3 Grades for energy rating system in HVAC in the subway station

诊断,评价地铁环控系统的用能特性,以验证用能诊断方法的可行性。收集某一时期该地铁站的运行数据,计算能耗等级并评价用能特性。表 2 为某一时期该地铁站不同用能模式下的能耗等级分布 P。

表 2 某一时期该地铁站不同用能模式下的能耗等级分布

Tab.2 Distribution table of the operational ratings of different modes in HVAC in the subway station

模式	$P_A/\%$	$P_B/\%$	$P_{C1}/\%$	$P_{C2}/\%$	$P_D/\%$	$P_E/\%$
1	0	0	0	0	0	100
2	11.65	62.13	0.01	0.01	0.03	20.39
3	100	0	0	0	0	0
4	78.92	18.07	3.01	0	0	0

在模式 1 中,所有实测耗能数据比该模式下的用能基准高。主要原因是:用能模式 1 主要包括地铁站早上与深夜的用能情况,此时地铁站环控系统的所有设备应当停止运行,但由于管理行为不规范,站台的进风机、排风机一直保持运行,导致当前条件下地铁站用能水平高于用能基准。

在模式 2 中,88% 实际耗能值低于用能基准,而 12% 的数据值高于基准。主要包括早上客流高峰前(7 点之前)和下午客流高峰后(18 点之后)的用能情况,高峰期客流量的不同导致冷机负荷发生波动,用能水平在用能基准值附近波动,而采样时期的室外环境温度降低,大部分情况下地铁站用能低于基准值。

模式 3 主要包括中午室外环境温度较高时的用能情况,采样时期的地铁站环控系统实际用能水平全部低于用能基准,主要原因是该时期室外气温降低,环控系统用能水平整体降低。

模式 4 主要包括早上和下午客流高峰时期的用

能情况,实际耗能数据全部低于用能基准,主要原因是采样时期的室外气温降低,环控系统用能水平整体降低。

综上所述,通过地铁站实际运行数据划分不同的用能模式,建立用能基准,并根据不同模式的用能基准对地铁站未来的用能进行诊断是可行的。此外,基于决策树模型划分的用能模式符合专业逻辑以及实际运行状况,用能诊断能够公正、准确地判断未来运行用能等级,并识别异常的用能模式,有助于查找地铁站运营漏洞、故障排查、优化运营管理模式,为地铁站节能工作的开展提供参考依据。

4 结论

本文提出了一种地铁站用能诊断方法,通过数据挖掘技术建立评价模型来评价地铁环控系统的用能特性。首先,通过相关性分析,确定影响地铁环控系统能耗的关键变量:室外环境温度、客流量;其次,根据所选取的关键变量,构建决策树模型来划分不同的用能模式,根据各个模式中能耗中位值作为每个模式的用能基准;最后,根据不同模式的用能基准对实际运行数据进行用能诊断,得到如下结论:

1) 地铁站的客流与室外温度对于地铁站的用能影响较大,采用决策树模型可以根据提取出的关键变量,科学、公正地划分地铁站不同用能模式,且符合专业逻辑以及实际运行状况。进而建立不同模式下的用能基准,实现地铁站用能诊断和异常能耗识别。

2) 实际数据的诊断结果表明:环境和客流的变化会引起用能水平的波动,但仍然贴近用能基准。一方面,该用能诊断方法能够识别异常用能模式,有助于查找地铁站运营漏洞、故障排查;另一方面,该用能诊断方法能够识别低能耗的用能模式,有利于优化环控模式,为运营管理提供决策。

本文受华中科技大学自主创新研究基金(5003120005)项目资助。(The project was supported by Independent Innovation Research Foundation of Huazhong University of Science and Technology(No. 5003120005).)

参考文献

[1] 常莉,冯炼,李鹏. 地铁环控系统不同区域能耗分析[J]. 制冷与空调(四川), 2009, 23(5): 115-118. (CHANG Li, FENG Lian, LI Peng. Energy consumption analysis of different areas on subway ECS[J]. Refrigeration and Air Conditioning, 2009, 23(5): 115-118.)

[2] WANG Yongcai, FENG Haoran, XI Xiangyu. Sense, model and identify the load signatures of HVAC systems in metro stations[R]. National Harbor, MD, United States: 2014 IEEE Power and Energy Society General Meeting, 2014.

[3] WANG Yongcai, FENG Haoran, QI Xiao. SEED: public energy and environment dataset for optimizing HVAC operation in subway stations[R]. National Harbor, MD, United States: 2014 IEEE Power and Energy Society General Meeting, 2014.

[4] ZHU Yingxin, QIN Ouyang. Airflow fluctuations and thermal environment: A literature review[R]. Beijing: The 4th International Symposium on Heating, Ventilating and Air Conditioning, 2003: 78-86.

[5] 朱颖心,秦绪忠,江亿. 站台屏蔽门在地铁热环境控制中的经济性分析[J]. 建筑科学, 1997(4): 21-25. (ZHU Yingxin, QIN Xuzhong, JIANG Yi. Economic analysis for platform screen door in subway thermal environment control[J]. Building Science, 1997(4): 21-25.)

[6] 江亿. 从实际运行数据出发开展地铁站节能工作[R]. 北京:清华大学, 2015. (JIANG Yi. Carrying out energy-saving work in subway stations based on actual operating data[R]. Beijing: Tsinghua University, 2015.)

[7] 石静雅,苏永清,岳继光. 轨道交通能耗影响因素分析及能耗评价体系的建立[J]. 铁道运输与经济, 2008, 30(9): 46-49. (SHI Jingya, SU Yongqing, YUE Jiguang. The analysis of influence factors in energy consumption of railway transport and the establishment of the energy evaluation system[J]. Railway Transport and Economy, 2008, 30(9): 46-49.)

[8] 孙宁,李照星,戴华明. 城市轨道交通能耗指标体系与节能措施研究——设备监理在节能减排中的创新实践[J]. 设备监理, 2011(1): 15-22. (SUN Ning, LI Zhaoxing, DAI Huaming. The study of energy consumption index system and energy saving measure of urban rail transit: the innovation practice of the equipment supervision in energy-saving emission reduction[J]. Plant Engineering Consultants, 2011(1): 15-22.)

[9] 郑玲芳. 上海市轨道交通规划环评能耗指标体系初探[J]. 环境科学与技术, 2013, 36(12): 402-405. (ZHENG Lingfang. The energy index system of Shanghai urban transit rail in strategy environmental assessment[J]. Environmental Science & Technology, 2013, 36(12): 402-405.)

[10] 陈焕新,刘江岩,胡云鹏,等. 大数据在空调领域的应用[J]. 制冷学报, 2015, 36(4): 16-22. (CHEN Huanxin, LIU Jiangyan, HU Yunpeng, et al. Application of big data in air-conditioning field[J]. Journal of Refrigeration, 2015, 36(4): 16-22.)

[11] QUINLAN J R. Induction of decision trees[J]. Machine

- learning, 1986, 1(1): 81-106.
- [12] 谢姐姐, 刘於勋. 决策树属性选择标准的改进[J]. 计算机工程与应用, 2010, 46(34): 115-118. (XIE Niuniu, LIU Yuxun. Improvement of attribute selection criterion of decision trees[J]. Computer Engineering and Applications, 2010, 46(34): 115-118.)
- [13] QUINLAN J R. C4.5: programs for machine learning [M]. Elsevier, 2014.
- [14] BREIMAN L, FRIEDMAN J H, OLSHEN R A, et al. Classification and regression trees[M]. CRC Press, 1984.
- [15] MEHTA M, AGRAWAL R, RISSANEN J. SLIQ: A fast scalable classifier for data mining[C]// International Conference on Extending Database Technology. Springer Berlin Heidelberg, 1996.
- [16] PARK H S, LEE M, KANG H, et al. Development of a new energy benchmark for improving the operational rating

system of office buildings using various data-mining techniques[J]. Applied Energy, 2016, 173: 225-237.

通信作者简介

陈焕新, 男, 教授, 华中科技大学制冷及低温工程系, (027) 87558330, E-mail: chenhuanxin@tsinghua.org.cn。研究方向: 制冷空调系统计算机模拟及优化, 制冷空调设备开发及新技术, 车辆制冷及其测控技术。

About the corresponding author

Chen Huanxin, male, professor, Department of Refrigeration & Cryogenics, Huazhong University of Science and Technology, +86 27-87558330, E-mail: chenhuanxin@tsinghua.org.cn. Research fields: computer simulation and optimization of refrigeration and air conditioning system, development and new technologies of refrigeration and air conditioning equipment, measurement and control technology of vehicle refrigeration.